AMENDMENTS TO THE CLAIMS:

Please cancel claims 12 to 53, and 68 to 98, without prejudice or disclaimer of subject matter thereof, amend claim 3, 4, 7, 8, 11, 54, 55 and 99, and add new claims 102 to 128, as shown below. This listing of claims replaces all prior versions, and listings, of claims in the application:

Listing of Claims:

1. to 2. (Cancelled)

3. (Currently Amended) The method of claim 99 further ~~including~~ comprising:

recognizing a gesture associated with the object by analyzing changes in the position information of the object, and

controlling the computer application based on the recognized gesture.

4. (Currently Amended) The method of claim 3 further ~~including~~ comprising:

determining an application state of the computer application; and

using the application state in recognizing the gesture.

5. (Previously Presented) The method of claim 99 wherein the object is the user.

6. (Previously Presented) The method of claim 99 wherein the object is a part of the user.

7. (Currently Amended) The method of ~~claim 99~~ claim 5 further ~~including~~ comprising providing feedback to the user relative to the computer application.

8. (Currently Amended) The method of claim 99 further comprising ~~including~~ mapping the position information from position coordinates associated with the object to screen coordinates associated with the computer application.

9 to 10. (Cancelled)

11. (Currently Amended) The method of claim 99 further ~~including~~ comprising:

analyzing the scene description to identify a change in position of the object; and

mapping the change in position of the object.

12 to 53. (Cancelled)

54. (Currently Amended) A stereo vision system for interfacing with an application program running on a computer, the stereo vision system comprising:

first and second video cameras arranged in an adjacent configuration and operable to produce ~~a series of~~ at least first and second stereo video images; and

a processor operable to receive the ~~series of~~ first and second stereo video images and detect objects appearing in an intersecting field of view of the cameras, the processor executing a process to:

define an object detection region in three-dimensional coordinates relative to a position of the first and second video cameras;

divide the first and second stereo video images into features;

pair features of the first stereo video image with features of the second stereo video image;

generate a depth description map, the depth description map describing the position and disparity of paired features relative to the first and second stereo video images;

generate a scene description based upon the depth description map, the scene description defining a three-dimensional position for each feature;

cluster adjacent features;

crop clustered feature based upon predefined thresholds;

analyze the three-dimensional position of each clustered feature within the object detection region to determine position information of a control object; and

~~select a control object as a cluster of features from the scene description appearing within the object detection region; and~~

map the position ~~coordinates~~ information of the control object to a position indicator associated with ~~the~~ an application program as the control object moves within the object detection region.

55. (Currently Amended) The stereo vision system of claim 54 wherein the process selects as ~~a control~~ the control object a detected object appearing closest to the video cameras and within the object detection region.

56. (Original) The stereo vision system of claim 54 wherein the control object is a human hand.

57. (Original) The stereo vision system of claim 54 wherein a horizontal position of the control object relative to the video cameras is mapped to an x-axis screen coordinate of the position indicator.

58. (Original) The stereo vision system of claim 54 wherein a vertical position of the control object relative to the video cameras is mapped to a y-axis screen coordinate of the position indicator.

59. (Original) The stereo vision system of claim 54 wherein the processor is configured to:

map a horizontal position of the control object relative to the video cameras to a x-axis screen coordinate of the position indicator;

map a vertical position of the control object relative to the video cameras to a y-axis screen coordinate of the position indicator; and

emulate a mouse function using the combined x-axis and y-axis screen coordinates provided to the application program.

60. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse using gestures derived from the motion of the object position.

61. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse based upon a sustained position of the control object in any position within the object detection region for a predetermined time period.

62. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse based upon a position of the position indicator being sustained within the bounds of an interactive display region for a predetermined time period.

63. (Original) The stereo vision system of claim 54 wherein the processor is further configured to map a z-axis depth position of the control object relative to the video cameras to a virtual z-axis screen coordinate of the position indicator.

64. (Original) The stereo vision system of claim 54 wherein the processor is further configured to:

map a x-axis position of the control object relative to the video cameras to an x-axis screen coordinate of the position indicator;

map a y-axis position of the control object relative to the video cameras to a y-axis screen coordinate of the position indicator; and

map a z-axis depth position of the control object relative to the video cameras to a virtual z-axis screen coordinate of the position indicator.

65 (Original) The stereo vision system of claim 64 wherein a position of the position indicator being within the bounds of an interactive display region triggers an action within the application program.

66. (Original) The stereo vision system of claim 54 wherein movement of the control object along a z-axis depth position that covers a predetermined distance within a predetermined time period triggers a selection action within the application program.

67. (Original) The stereo vision system of claim 54 wherein a position of the control object being sustained in any position within the object detection region for a predetermined time period triggers a selection action within the application program.

68. to 98. (Cancelled).

99. (Currently Amended) A method of using computer vision to interface with a computer, the method comprising:

capturing at least first and second images of a scene;

dividing the first and second images into features;

pairing features of the first image with features of the second image;

generating a depth description map, the depth description map describing the position and disparity of paired features relative to the first and second images;

generating a scene description based upon the depth description map, the scene description defining ~~that includes an indication of~~ a three-dimensional position for each ~~of a~~ feature ~~included in a scene~~;

clustering adjacent features;

cropping clustered feature based upon predefined thresholds;

defining an object detection region;

analyzing the ~~scene description including the indication of the~~ three-dimensional position of each clustered ~~the~~ feature within the object detection region to determine position information of an object ~~within the scene~~; and

using the position information to control a computer application.

100. (Previously Presented) The method of claim 99 wherein generating the scene description comprises generating the scene description from stereo images.

101. (Previously Presented) The method of claim 99 wherein:

generating a scene description comprises generating a scene description that includes an indication of a three-dimensional position of a feature included in a scene and an indication a shape of the feature; and

analyzing the scene description comprises analyzing the scene description including the indication of the three-dimensional position of the feature and the indication of the shape of the feature to determine position information of an object.

102. (New) A method for video-based control of an application program, comprising the steps of:

defining a region of interest, wherein the region of interest is within a field of view of an image detector;

acquiring at least one image of the region of interest and a scene surrounding the region of interest;

producing a scene description based upon the at least one image;

defining an object detection region within the region of interest based upon the scene description;

measuring a position of an object within the object detection region;

mapping the position of the object as a representation in the application program; and

displaying the representation.

103. (New) The method of claim 102 further comprising the steps of:

measuring a change in the position of the object;

interpreting the change as a gesture;

mapping the gesture to the representation; and

controlling the application program with the representation.

104. (New) The method of claim 102 further comprising the step of performing a stereo image analysis on the at least one image.

105. (New) The method of claim 102 wherein the object is a human hand.

106. (New) The method of claim 102 wherein the position is expressed in a world coordinate system.

107. (New) The method of claim 102 wherein the position is expressed in an X-Y-Z coordinate system.

108. (New) The method of claim 102 wherein the region of interest is a three-dimensional region of interest.

109. (New) The method of claim 102 wherein the object detection region is a three-dimensional object detection region.

110 (New) The method of claim 103 wherein controlling the application program further comprises moving a cursor.

111. (New) The method of claim 103 wherein controlling the application program further comprises selecting a control.

112. (New) The method of claim 103 wherein interpreting the change as a gesture is context-sensitive.

113. (New) The method of claim 102 wherein defining the object detection region is based upon expected location of the object within the scene description.

114. (New) The method of claim 102 wherein defining the object detection region is based upon shape of the object within the scene description.

115. (New) The method of claim 102 wherein defining the object detection region is based upon pose of the object within the scene description.

116. (New) The method of claim 102 wherein defining the object detection region is based upon an anatomical model.

117. (New) The method of claim 102 wherein producing the scene description further comprises the step of producing a background reference.

118. (New) The method of claim 117 wherein producing the scene description further comprises the step of cropping the background reference.

119 (New) The method of claim 118 wherein producing the scene description further comprises the step of clustering adjacent features in at least one image, based upon predefined criteria.

120. (New) The method of claim 119 wherein defining the object detection region further comprises the step of determining object presence based upon the clustered features.

121. (New) The method of claim 102 wherein the scene description is a three-dimensional scene description.

122. (New) A system comprising:

an image detector;

a display; and

a processor, said processor executing an application program and a process to:

define a region of interest, wherein the region of interest is within a field of view of the image detector,

acquire at least one image of the region of interest and a scene surrounding the region of interest,

produce a scene description based upon the at least one image,

define an object detection region within the region of interest based upon the scene description,

measure a position of an object within the object detection region,

map the position of the object as a representation in an application program, and

display the representation.

123. (New) The system of claim 122 wherein said processor further executes an application program and a process to:

measure a change in the position of the object;

interpret the change as a gesture;

map the gesture to the representation; and

control the application program with the representation.

124. (New) The system of claim 122 wherein the image detector is a stereo vision detector.

125. (New) The system of claim 122 wherein the image detector is a video camera.

126. (New) The system of claim 122 wherein the application program is a graphical user interface ("GUI").

127. (New) The system of claim 122 wherein the application program is a video game.

128. (New) The system of claim 122 where the image detector is an overhead image detector.